# Modelling European Industrial Production with Multivariate Singular Spectrum Analysis: A Cross Industry Analysis*

Emmanuel Sirimal Silva[§], Hossein Hassani[†], and Saeed Heravi[‡]
[§]*Fashion Business School, London College of Fashion, University of the Arts London, UK*
[†]*Research Institute of Energy Management and Planning, University of Tehran, Tehran, Iran*
[‡]*Cardiff Business School, Cardiff University, UK*

December 13, 2017

**Abstract**

In this paper, an optimized Multivariate Singular Spectrum Analysis (MSSA) approach is proposed to find leading indicators of cross industry relations between 24 monthly, seasonally unadjusted industrial production (IP) series for German, French and UK economies. Both Recurrent and Vector forecasting algorithms of Horizontal MSSA (HMSSA) are considered. The results from the proposed multivariate approach are compared with those obtained via the optimized univariate Singular Spectrum Analysis (SSA) forecasting algorithm to determine the statistical significance of each outcome. The data is rigorously tested for normality, seasonal unit root hypothesis and structural breaks. The results are presented such that users can not only identify the most appropriate model based on the aim of the analysis, but also easily identify the leading indicators for each IP variable in each country. Our findings show that for all three countries, forecasts from the proposed MSSA algorithm outperform the optimized SSA algorithm in over 70% of the cases. Accordingly, this new approach succeeds in identifying leading indicators and is a viable option for selecting the SSA choices $L$ and $r$ which minimises a loss function.

Keywords: Forecasting; Singular Spectrum Analysis; Multivariate SSA; industrial production; leading indicators.

## 1 Introduction

The introduction of the nonparametric time series analysis and forecasting technique of Singular Spectrum Analysis (SSA) is closely associated with the work of Broomhead and King (1986a,b). Since then, SSA has progressed rapidly and transformed itself into a powerful technique which is increasingly exploited for providing solutions to real world problems in a variety of different fields, see for example, Gong et al. (2017), Merte (2017), Yu et al. (2017), Mahmoudvand and Rodrigues (2017), Khan and Poskitt (2017), Hassani et al. (2017; 2016; 2015), Lai and Guo (2017), Ghodsi et al. (2015), and Silva and Hassani (2015). Few reasons underlying this augmented usage of SSA can be partly attributed to its nonparametric nature. This implies that the parametric assumptions relating to normality, stationarity and linearity do not apply when

*Article accepted for publication in the Journal of Forecasting, DOI:10.1002/for.2508

modelling with SSA. In addition, SSA is popular for its sound filtering capabilities which enables signal extraction from any given time series and facilitates a richer analysis (see for example, Ghodsi et al. (2015)). Furthermore, over the years, SSA has been successful in analysing and forecasting both stationary and non-stationary time series which is of utmost importance to the financial and economic fields where uncertainty and recessions are of major concern.

As noted above, SSA (used to refer to the univariate version) is widely exploited at present and there exists a considerable number of applications. However, in comparison to SSA, Multivariate SSA (MSSA) is considered to be in its early stages. This is clearly visible should one pursue a literature search which would uncover only a handful of MSSA applications. Few of the recent examples of MSSA based applications are found in Silva et al. (2017), de Carvalho and Rua (2017), Hassani and Silva (2016), Portes and Aguirre (2016), and Hassani et al. (2013). The MSSA technique is based on two choices known as the window length $L$, and the number of eigenvalues $r$ (Sanei and Hassani, 2015). The selection of SSA and MSSA choices have for long been a point of discussion. Those interested in reading on the different approaches available are referred to Khan and Poskitt (2013) and Sanei and Hassani (2015). In brief, the approaches explained in the aforementioned publications are time consuming and imposes a restriction on the useability of SSA and MSSA by individuals who are not conversant with the methodology. Recently, in Hassani et al. (2015) an algorithm was presented for optimizing the univariate SSA approach and obtaining the optimized SSA choices based on a loss function. This meant that users would no longer require an in-depth knowledge on the theory underlying SSA and could apply it universally to any given forecasting problem.

In hope of encouraging more users to exploit the MSSA technique, through this paper we present an algorithm for determining the optimal MSSA choices by minimising a given loss function. In brief, this optimization procedure will indicate the best possible forecast attainable via MSSA by minimising any given loss function. In a forecasting exercise this would mean splitting the series into three parts; training, testing and validation whereby we would use the proposed algorithm on the training and testing data to ascertain the optimal MSSA choices and then evaluate its performance over the validation set. However, in this paper we are mainly concerned with exploiting this new algorithm for finding leading indicators for modelling and forecasting industrial production (IP) data. In simple terms, here we would consider splitting all the data into two parts; training and test sets alone and find the optimal MSSA choices for forecasting the test set. In the process, we would evaluate all possible dual combinations. Then, the forecast attainable via the resulting MSSA choices which provide the lowest level in a given loss function would be representative of the best possible MSSA forecast. As such, this optimization procedure which minimises a loss function will enable the identification of leading indicators for a given variable.

The feasibility of the proposed approach is evaluated via an application which considers modelling and forecasting 24 monthly, seasonally unadjusted industrial production indices for UK, France and Germany which have previously been analysed in both linear and nonlinear contexts in Hassani et al. (2013), Hassani et al. (2009), Heravi et al. (2004) and Osborn et al. (1999). The output from this study will be important to economists, policy makers and practitioners as they will be able to identify the leading indicators for each industrial production variable. This in turn helps better understand the underlying dynamics, saves time and improves the accuracy of forecasts attainable via multivariate modelling. It is noteworthy that in Hassani et al. (2013) the authors have used basic MSSA for forecasting UK industrial production and those results cannot be approximated as the optimal leading indicators for IP as there was no optimization procedure involved in the modelling process.

The significance and importance of these 24 series have been discussed in detail in the aforementioned publications and are therefore not reproduced here. In Osborn et al. (1999) it was found that these 24 time series are highly seasonal and that these seasonality levels are much larger than those reported for monthly output in the United States (Miron, 1996). However, in terms of linearity it was reported in Heravi et al. (2004) that there was no substantial evidence of non-linearity in majority of the cases. Yet, it should be noted that these conclusions were based on data ending in December 1995.

In contrast, we consider more up-to-date observations which end in 2014 and in turn we are able to provide a more timely analysis. The forecasts from the proposed optimal MSSA forecasting approach are compared with forecasts from the optimal SSA approach in Hassani et al. (2015). We do not consider other models such as ARIMA, Holt-Winters and Vector Autoregressive models as in Hassani et al. (2013) and Hassani et al. (2009) it has already been shown that SSA and MSSA can outperform these approaches when modelling and forecasting European industrial production.

Accordingly, this paper has several key contributions. Firstly, this is the first study in which the optimized SSA and optimized MSSA approaches are evaluated for modelling and forecasting European industrial production. Secondly, we propose the use of the Horizontal MSSA (HMSSA) algorithm introduced in this work as a method for finding leading indicators. Thirdly, the study considers both vector and recurrent forecasting algorithms of SSA giving the reader a complete outlook into the performance SSA as a whole when forecasting IP.

The remainder of this manuscript is organized such that: Section 2 presents the optimized MSSA forecasting algorithm; Section 3 introduces the data and provides a thorough analysis; Section 4 reports the empirical results from our application and Section 5 concludes the paper.

# 2 Methodology

## 2.1 Optimized SSA Forecasting Algorithm

The optimal SSA forecasting algorithm is documented in Hassani et al. (2015) and those interested are referred to that paper.

## 2.2 Optimized HMSSA Forecasting Algorithm

In what follows we present the optimized HMSSA forecasting algorithms for finding leading indicators of European industrial production. Those interested in an in-depth explanation of the theory underlying MSSA are directed to Sanei and Hassani (2015). We begin by considering the HMSSA-R optimal forecasting algorithm which is followed by the HMSSA-V optimal forecasting algorithm. In presenting these two algorithms we mainly follow and rely on the notations in Sanei and Hassani (2015).

### 2.2.1 HMSSA-R Optimal Forecasting Algorithm

1. Consider $M$ time series with identical series lengths of $N_i$, such that $Y_{N_i}^{(i)} = (y_1^{(i)}, \dots, y_{N_i}^{(i)})$ $(i = 1, \dots, M)$.

2. For forecasting exercises we would split each time series into three parts leaving $\frac{2}{3}^{rd}$ for model training and testing, and $\frac{1}{3}^{rd}$ for validation. However, as the objective here is to exploit this same algorithm to find leading indicators we split the time series into training and test sets alone.

3. Beginning with a fixed value of $L = 2$ ($2 \leq L \leq \frac{N}{2}$) and in the process, evaluating all possible values of $L$ for $Y_{N_i}$, using the training data construct the trajectory matrix $\mathbf{X}^{(i)} = [X_1^{(i)}, \ldots, X_K^{(i)}] = (x_{mn})_{m,n=1}^{L,K_i}$ for each single series $Y_{N_i}^{(i)}$ ($i = 1, \ldots, M$) separately.

4. Then, construct the block trajectory matrix $\mathbf{X}_H$ as follows:

$$\mathbf{X}_H = \left[ \ \mathbf{X}^{(1)} : \ \ \mathbf{X}^{(2)} : \ \ \cdots \ \ : \mathbf{X}^{(M)} \ \right].$$

5. Let vector $U_{H_j} = (u_{1j}, \ldots, u_{Lj})^T$, with length $L$, be the $j^{th}$ eigenvector of $\mathbf{X}_H \mathbf{X}_H^T$ which represents the SVD.

6. Evaluate all possible combinations of $r$ ($1 \leq r \leq L - 1$) step by step for the selected $L$ and construct $\widehat{\mathbf{X}}_H = \sum_{i=1}^{r} U_{H_i} U_{H_i}^T \mathbf{X}_H$ as the reconstructed matrix obtained using $r$ eigentriples:

$$\mathbf{X}_H = \left[ \ \widehat{\mathbf{X}}^{(1)} : \ \ \widehat{\mathbf{X}}^{(2)} : \ \ \cdots \ \ : \widehat{\mathbf{X}}^{(M)} \ \right].$$

7. Consider matrix $\widetilde{\mathbf{X}}^{(i)} = \mathcal{H}\widehat{\mathbf{X}}^{(i)}$ ($i = 1, \ldots, M$) as the result of the Hankelization procedure of the matrix $\widehat{\mathbf{X}}^{(i)}$ obtained from the previous step for each possible combination of SSA choices.

8. Let $U_{H_j}^{\triangledown}$ denote the vector of the first $L - 1$ coordinates of the eigenvectors $U_{H_j}$, and $\pi_{H_j}$ indicate the last coordinate of the eigenvectors $U_{H_j}$ ($j = 1, \ldots, r$).

9. Define $\upsilon^2 = \sum\limits_{j=1}^{r} \pi_{H_j}^2$.

10. Denote the linear coefficients vector $\mathcal{R}$ as follows:

$$\mathcal{R} = \frac{1}{1 - \upsilon^2} \sum_{j=1}^{r} \pi_{Hj} U_{Hj}^{\triangledown}. \tag{1}$$

11. If $\upsilon^2 < 1$, then the $h$-step ahead HMSSA forecasts exist and is calculated by the following formula:

$$\left[\hat{y}_{j_1}^{(1)}, \ldots, \hat{y}_{j_M}^{(M)}\right]^T = \begin{cases} \left[\tilde{y}_{j_1}^{(1)}, \ldots, \tilde{y}_{j_M}^{(M)}\right], & j_i = 1, \ldots, N_i, \\ \mathcal{R}^T \mathbf{Z}_h, & j_i = N_i + 1, \ldots, N_i + h, \end{cases} \tag{2}$$

where, $\mathbf{Z}_h = \left[Z_h^{(1)}, \ldots, Z_h^{(M)}\right]^T$ and $Z_h^{(i)} = \left[\hat{y}_{N_i-L+h+1}^{(i)}, \ldots, \hat{y}_{N_i+h-1}^{(i)}\right]$ ($i = 1, \ldots, M$).

12. Seek the combination of $L$ and $r$ which minimises a loss function, $\mathcal{L}$ and thus represents the optimal HMSSA-R choices for decomposing and reconstructing in a multivariate framework.

13. Finally use the selected optimal $L$ to decompose the series comprising of the validation set, and then select $r$ singular values for reconstructing the less noisy time series, and use this newly reconstructed series for forecasting the remaining $\frac{1}{3}^{rd}$ observations (or the test set as relevant to this study).

#### 2.2.2 HMSSA-V Optimal Forecasting Algorithm

1. Begin by following the steps in 1-9 of the HMSSA-R optimal forecasting algorithm above.

2. Consider the following matrix

$$\mathbf{\Pi} = \mathbf{U}^{\nabla}\mathbf{U}^{\nabla T} + (1-v^2)RR^T, \tag{3}$$

   where $\mathbf{U}^{\nabla} = [U_1^{\nabla}, ..., U_r^{\nabla}]$. Now consider the linear operator

$$\mathcal{P}^{(v)} : \mathfrak{L}_r \mapsto \mathbb{R}^L, \tag{4}$$

   where

$$\mathcal{P}^{(v)}Y = \begin{pmatrix} \mathbf{\Pi}Y_{\vartriangle} \\ R^T Y_{\vartriangle} \end{pmatrix}, \; Y \in \mathfrak{L}_r, \tag{5}$$

   and $Y_{\vartriangle}$ is vector of last $L-1$ elements of $Y$.

3. Define vector $Z_j^{(i)}$ $(i = 1, \ldots, M)$ as follows:

$$Z_j^{(i)} = \begin{cases} \widetilde{X}_j^{(i)} & \text{for } j = 1, \ldots, k_i \\ \mathcal{P}^{(v)} Z_{j-1}^{(i)} & \text{for } j = k_i+1, \ldots, k_i+h+L-1 \end{cases} \tag{6}$$

   where, $\widetilde{X}_j^{(i)}$'s are the reconstructed columns of trajectory matrix of the $i^{\text{th}}$ series after grouping and leaving noise components.

4. Now, by constructing matrix $\mathbf{Z}^{(i)} = [Z_1^{(i)}, ..., Z_{k_i+h+L-1}^{(i)}]$ and performing diagonal averaging we obtain a new series $\hat{y}_1^{(i)}, ..., \hat{y}_{N_i+h+L-1}^{(i)}$, where $\hat{y}_{N_i+1}^{(i)}, ..., \hat{y}_{N_i+h}^{(i)}$ provides the $h$ step ahead HMSSA-V forecast for the selected combination of $L$ and $r$.

5. Finally, follow steps 12-13 in the HMSSA-R optimal forecasting algorithm to find the optimal $L$ and $r$ for obtaining HMSSA-V forecasts.

# 3 Data

The data used in this study are taken from I.N.S.E.E (Institute National de la Statistiuqe et des Etudes Economiques) for France, from Statistisches Bundesamt, Wesbaden for Germany and from the Office for National Statistics (ONS) for the UK. It includes eight major components of real industrial production in France, Germany and the UK. The series are seasonally unadjusted monthly indices for real output in Electricity/Gas, Chemicals, Fabricated Metals, Vehicle, Food Products, Basic Metals, Electrical Machinery and Machinery. The eight series considered in this study are important, diverse and account for more than 50% of the total industrial production in each country. The same eight industries have been considered by Hassani et al. (2009), Heravi et al. (2004) and Osborn (1999).

In all cases, our sample period ends in February 2014. However, the data for France start from January 1990, for Germany from 1991 and for the UK from January 1998. Figures 1-3 show the time series used in this study. As visible, most series for France and Germany present a long period of growth in 1990s and up to the current recession of 2008/2009. For the UK, however, most series show a period of stagnation in early 2000s and following the recession in 2008. For almost all the series, the impact of current recession which is attributed to banking crisis can be seen in 2008. The only exceptions are Food Products and Electricity/Gas production which are inelastic in nature. It is also evident that seasonality is the dominant pattern in the data and the use of a filtering technique such as SSA/MSSA in such cases could provide better results as they are able to filter, extract and model such signals.

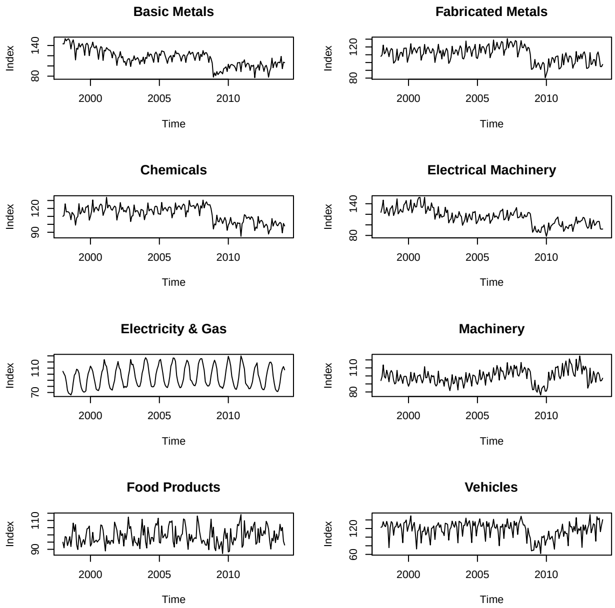


Figure 1: UK industrial production indices.

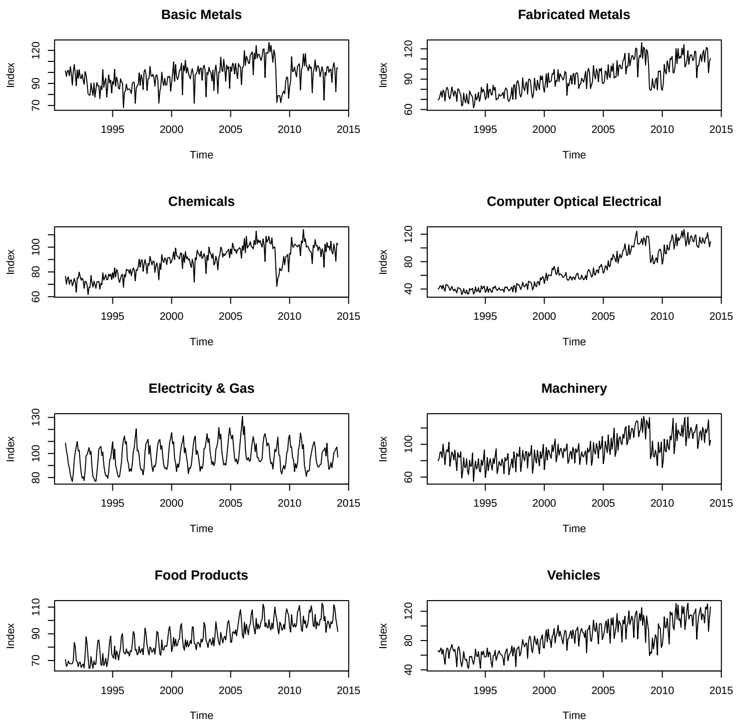


Figure 2: German industrial production indices.

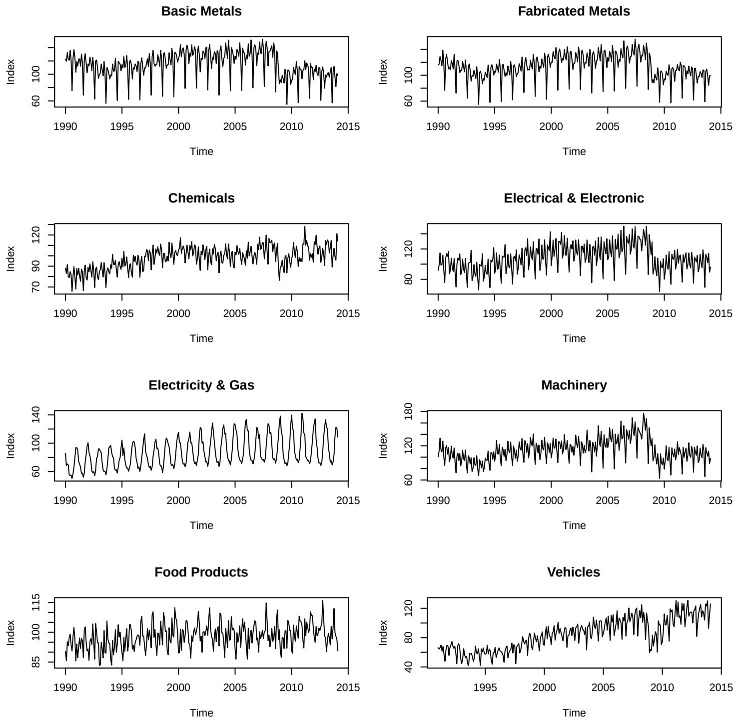


Figure 3: France industrial production indices.

Table 1 shows the weights and descriptive statistics for the yearly percentage changes in the original series, i.e. $100((Y_t - Y_{t-12})/(Y_{t-12})$. For Germany, all sectors have experienced growths over the whole period. In particular, Electrical Machinery and Vehicles in Germany show substantial growth in production with average increases of around 4.0 and 2.5 percentages per year. In contrast, the results for the UK are all negative, with the exception of Vehicles, Electricity/Gas and Food Products. Basic Metals have experienced the highest decline in the UK with an average of about 2%. For France, the highest growth was reported for Electricity/Gas and Basic Metals had the highest decline over the sample period. The sample standard deviations indicate greater volatility for the Vehicles series than those of other sectors, with very low volatility for Food Products. The results for normality testing based on the Shapiro-Wilk (SW) test (Royston, 1982a,b; Royston, 1995) also provide strong evidence of non-normality for all the series, except for the Electricity/Gas, Electrical Machinery and Food Products. The last column in Table 1 presents the OCSB test results (Osborn et al., 1988) for seasonal unit hypothesis. The results indicate minimal evidence for non-stationary seasonal unit roots in these monthly industrial production data, except within some traditional industrial sectors such as Vehicles, Basic Metals, Machinery, and Fabricated Metals.

Table 1: Descriptive statistics for UK, German and France IP Indices.

| | Weight | Mean | SD | SW($p$) | OCSB |
|---|---|---|---|---|---|
| *UK* | | | | | |
| Basic Metals | 3.00 | -1.98 | 13.50 | <0.01 | 0 |
| Fabricated Metals | 3.80 | -0.63 | 9.71 | <0.01 | 1 |
| Chemicals | 8.50 | -0.76 | 8.27 | <0.01 | 0 |
| Electrical Machinery | 4.70 | -1.44 | 12.15 | <0.01 | 0 |
| Electricity & Gas | 10.20 | 0.50 | 10.35 | 0.22* | 0 |
| Machinery | 6.70 | -0.16 | 12.44 | <0.01 | 1 |
| Food Products | 7.50 | 0.15 | 7.01 | 0.14* | 0 |
| Vehicles | 5.80 | 0.19 | 24.07 | <0.01 | 1 |
| *Germany* | | | | | |
| Basic Metals | 5.60 | 0.04 | 14.78 | <0.01 | 0 |
| Fabricated Metals | 4.50 | 1.64 | 11.91 | <0.01 | 0 |
| Chemicals | 8.60 | 1.27 | 9.95 | <0.01 | 0 |
| Computer Optical Electrical | 10.40 | 3.95 | 14.57 | 0.10* | 0 |
| Electricity & Gas | 7.60 | 0.25 | 7.69 | 0.20* | 0 |
| Machinery | 6.50 | 1.08 | 17.91 | 0.09* | 0 |
| Food Products | 13.60 | 1.46 | 6.78 | 0.08* | 0 |
| Vehicles | 7.20 | 2.45 | 19.94 | <0.01 | 0 |
| *France* | | | | | |
| Basic Metals | 3.90 | -0.92 | 27.23 | <0.01 | 1 |
| Fabricated Metals | 4.30 | -0.79 | 26.21 | <0.01 | 1 |
| Chemicals | 8.90 | 1.01 | 11.59 | 0.02* | 0 |
| Electrical & Electronic | 7.10 | 0.05 | 21.59 | <0.01 | 1 |
| Electricity & Gas | 9.00 | 1.42 | 13.03 | 0.07* | 0 |
| Machinery | 9.60 | -0.29 | 22.74 | <0.01 | 1 |
| Food Products | 8.60 | 0.20 | 6.87 | <0.01 | 0 |
| Vehicles | 9.80 | 0.15 | 45.81 | <0.01 | 0 |

*Note*: * indicates the data is normally distributed based on the Shapiro-Wilk (SW) test (Royston, 1982a,b; Royston, 1995) at $p = 0.05$.
0 indicates there is no seasonal unit root based on the OCSB test (Osborn et al., 1988) at $p$=0.05.
1 indicates there is a seasonal unit root based on the OCSB test (Osborn et al., 1988) at $p$=0.05.

Table 2 shows the Bai and Perron (2003) test output for structural breaks in these time series. Almost all the series show a break or multiple breaks in the data period considered, except the Food Products in UK. Results for all the three countries, based on the Bai Perron test, indicate that almost all sectors are affected by the current recession of 2008, except for the Electricity/Gas and Food Products sectors.

Table 2: Breakpoints for all major components of the UK, German and France IP Indices.

| | UK | Germany | France |
|---|---|---|---|
| BM | 2001(7) 2004(2) 2008(10) | 2000(1) 2005(5) 2008(10) | 1999(8) 2008(7) |
| FM | 2005(2) 2008(11) 2011(4) | 1997(3) 2000(8) 2005(5) 2010(8) | 1999(8) 2008(7) |
| C | 2006(2) 2008(10) | 1996(2) 1999(8) 2004(12) 2008(10) | 1993(12) 1997(8) |
| EM | 2001(11) 2008(12) 2011(8) | 1996(2) 1999(8) 2005(8) 2010(8) | 1997(8) 2008(12) |
| E&G | 2000(9) | 1995(10) | 1998(9) |
| M | 2006(2) 2008(9) 2011(2) | 1999(8) 2005(8) | 1995(2) 2004(8) 2008(12) |
| FP | None | 1994(8) 1999(8) 2005(7) | 1995(8) |
| V | 2008(7) 2010(12) | 998(1) 2002(2) 2010(8) | 1998(8) 2008(7) |

# 4 Empirical Results

We now turn to an issue of central interest in this paper; finding leading indicators for modelling industrial production data and the evaluation of forecast performance between the optimized SSA and MSSA algorithms. In the process, we assess the forecasting accuracy of Recurrent SSA (SSA-R) and Vector SSA (SSA-V) with HMSSA-R and HMSSA-V for four different horizons, 1-step ahead, 3 and 6-steps ahead and one year ahead. All models are estimated based on the data up to June 2011 as in-sample for training and post-sample forecasts are then computed for the last 32 months ending February 2014.

Forecast accuracy is measured based on the magnitude of forecast errors from the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) criteria. These measures reported quantitatively similar results. Therefore, to save space, we only report the RMSE, as this the is the most frequently quoted measure in forecasting (Altavilla and De Grauwe, 2010; Sadorsky, 2016; Hirashima et al., 2017). Tables 3-5 show the out-of-sample RMSE and Relative RMSE (RRMSE) results for UK, Germany and France. The ratio of RMSE is defined as:

$$\text{RRMSE} = \frac{MSSA}{SSA} = \frac{\left(\sum_{i=1}^{N}(\widehat{y}_{T+h,i} - y_{T+h,i})^2\right)^{1/2}}{\left(\sum_{i=1}^{N}(\widetilde{y}_{T+h,i} - y_{T+h,i})^2\right)^{1/2}},$$

where, $\widehat{y}_{T+h}$ is the $h$-step ahead forecast obtained by MSSA, $\widetilde{y}_{T+h}$ is the $h$-step ahead forecast from the SSA model, and $N$ is the number of the forecasts. If $\frac{MSSA}{SSA}$ is less than 1, then MSSA outperforms SSA by 1-$\frac{MSSA}{SSA}$ percent and vice versa.

## 4.1 Cross Industry Analysis

Tables 3-5 report the out-of-sample forecasting results for the IP for each country. Forecasts are calculated at horizons of 1, 3, 6 and 12 months ahead such that both the long and short run variation is accounted for. When reporting these results, we have considered the indicators at lag 0 and lag 12, but to save space we only report the result with the lag which produced the lower RMSE. Note that in addition to reporting the RMSE and RRMSE, shown at the bottom of each table are some interesting summary statistics based on the score and percentage. Another important point is that the multivariate forecasting indicators which are reported here are the best possible outcomes for MSSA. These have been determined by evaluating each component of IP with every other component of IP for each respective country and selecting the combination of indicators which gives the lowest forecasting error. Accordingly, the combinations shown below are the leading cross industry indicators for industrial production components in UK, Germany and France[1].

We begin by considering the results for UK. The first point to note is that the results have been presented in a manner which enables readers to identify the best model for a particular component given their horizon of interest. The general observation is that MSSA forecasts are

[1]We also went a step further to evaluate the cross country relations to see whether it might be possible to identify leading indicators across countries. On this occasion we specifically sought to model production in the same industry across the different countries. This meant for example evaluating whether the Germany and France Basic Metals indices would act as a leading indicator for the UK Basic Metals. This approach did not yield any positive results. As such we do not report the results in this paper, but they are available upon request. It should be noted that this does not imply there are no cross country relations, but simply that one should look for different combinations which consider different industries. Such an effort could yield positive outcomes but it remains beyond the mandate and aim of this paper.

able to outperform SSA forecasts 81% of the time for all components and horizons considered here. The HMSSA-R forecasts have outperformed SSA-R forecasts 84% of the time whilst HMSSA-V forecasts outperform SSA-V forecasts 78% of the time. The findings also show that HMSSA-V forecasts have outperformed HMSSA-R forecasts 56% of the time. These summary statistics clearly indicate that, for UK, more accurate forecasts can be obtained for the IP indices in majority of the cases by employing MSSA. Interestingly, we find that in the case of univariate SSA, both SSA-R and SSA-V models are indifferent as their forecasts outperform each other on an equal number of cases. Whilst the above conclusions are based on the number of times where forecasts from one model outperforms another based on the RMSE, the RRMSE criterion enables quantification of the superiority of each model in comparison to another.

The average RMSE values for all four models and the average RRMSE for comparing between the MSSA and SSA models are also reported in each of the tables below. The first point to note is that based on the average RRMSE, when forecasting all eight components of UK IP across the forecasting horizons of 1, 3, 6 and 12 steps ahead, on average the HMSSA-R forecasts are 8% better than SSA-R forecasts whilst the HMSSA-V forecasts are 6% better than SSA-V forecasts. The average RMSE column confirms this result and shows that HMSSA-R is the only model which reports an average error of less than 5 across all horizons and all components. Interestingly, SSA-R and SSA-V not only report equal performance based on the score but also perform equally in terms of the average RMSE as well. Finally, it is clear that for UK if we were interested in a single model for all components across all these horizons, then based on the score, total and overall average RMSE and RRMSE, the most appropriate model would be the HMSSA-R model with the corresponding multivariate indicator. Finally, based on our findings the leading indicators for UK IP components are shown within brackets in Table 3.

Next, we consider the out-of-sample forecasting results for all eight components of the IP for Germany which are reported in Table 4. In general, we can see once again that MSSA forecasts are seen outperforming SSA forecasts 81% of the time as was witnessed in the case of UK. If we compare the score between MSSA and SSA models, then forecasts from HMSSA-R have outperformed SSA-R forecasts 72% of the time whilst forecasts from HMSSA-V have outperformed SSA-V forecasts 78% of the time. In terms of which MSSA model has captured majority of the score, the summary statistics indicate that forecasts from HMSSA-R has only outperformed those from HMSSA-V 38% of the time whilst HMSSA-V forecasts have outperformed HMSSA-R forecasts 62% of the time. These results indicate that for Germany, based on the score it appears that HMSSA-V is the most appropriate single model for forecasting all eight components of IP. This is interesting as for UK we noticed HMSSA-R being recognized as the best model based on the score. At the same time, unlike what was experienced with UK, based on the score criterion we are able to suggest that SSA-V provides forecasts which outperforms SSA-R 72% of the time when forecasting the German IP components.

The total RMSE values indicate that HMSSA-V does in fact report the lowest overall error when forecasting all eight IP components across all horizons for Germany. However, the overall errors are almost identical between the two MSSA models and so in general they appear to be indifferent. Yet, in comparison to SSA there are positive gains with using MSSA as HMSSA-R forecasts are seen reporting accuracy levels which are 8% greater than SSA-R whilst in comparison to SSA-V, HMSSA-V forecasts are found to be 7% better in terms of accuracy as suggested by the overall average RRMSE values.

Once again we provide a more in-depth analysis by considering the variation in forecasting results based on the horizon of interest. If forecasting at $h = 1$ month ahead is the purpose, then based on the average lowest RMSE we can present the HMSSA-R model to be the most appropriate. At obtaining three-steps ahead forecasts, the HMSSA-V model is found to be the best average performer with a similar conclusion at $h = 6$ months ahead. For forecasting at $h = 12$ steps ahead the HMSSA-R model is once again seen reporting the lowest average RMSE and is thus the recommended model. For univariate forecasting, at $h = 1$, 3 and 6 steps ahead, the SSA-V model reports the lowest average RMSE whilst at $h = 12$ steps ahead the SSA-R model is likely to be more appropriate. Therefore, based on our results, it is clear that once again the optimized MSSA algorithm has outperformed the optimized SSA algorithm (as seen with UK). If we are interested in obtaining the most number of outcomes with the lowest forecasting error across all horizons, then we should rely on the HMSSA-V model with the combinations reported in Table 4, whilst for forecasting at certain horizons alone we should rely on different models as advised above. Once again, the presentation of results in Table 4 enables the reader to determine which model is best for each component of German IP at each horizon.

Finally, we consider the out-of-sample forecasting results for the IP in France. The results are reported in Table 5. In general, once again MSSA is seen outperforming SSA forecasts 81% of the time. HMSSA-R forecasts have outperformed SSA-R forecasts 84% of the time whilst HMSSA-V forecasts have outperformed SSA-V forecasts 78% of the time. In terms of the best MSSA model, based on the score, HMSSA-R takes precedence over HMSSA-V. Likewise, in terms of the best univariate model, based on the score, SSA-V takes precedence over SSA-R.

The total RMSE statistic also clearly indicates that HMSSA-R provides the lowest overall error across all horizons when forecasting all eight components of French IP. Furthermore, in comparison to SSA-R, SSA-V reports the lower total error. The overall average RRMSE criterion indicates that HMSSA-R forecasts across all components and all horizons are 11% better than SSA-R forecasts whilst HMSSA-V forecasts for the same conditions are seen reporting 7% better forecasts than SSA-V. The gain of 11% reported by HMSSA-R is the highest in comparison to the UK and France results.

Finally, we consider the variation in overall forecasting results based on the forecasting horizon. Firstly, we notice that at horizons of 1 step-ahead, HMSSA-V provides the overall lowest average RMSE whereas at 12 steps-ahead, both HMSSA-V and HMSSA-R are almost the same on average. In the medium term, i.e. at $h = 6, 12$ steps-ahead, HMSSA-R is seen reporting the overall lowest average error. In terms of univariate modelling, at $h = 1, 3$ *and* 12 months ahead, SSA-V reports the overall lowest average error, whilst at $h = 6$ months ahead, SSA-R is seen reporting the lowest error.

The presentation of results provide added value to this manuscript. This is because, based on the aim of the forecasting exercise, readers can differentiate and identify which SSA or MSSA model is best. The different options made available for selection of the most appropriate model are as follows. (1) The best model for any given component at any given forecasting horizon. (2) If obtaining the highest number of positive outcomes with the lowest forecasting error using a single model is important, then one can decide on such a model based on the score and percentage. (3) If the overall lowest forecasting error attainable using a single model is of importance given that one has to forecast all components across all horizons, then the option is available to decide on such a model by looking at the RMSE and RRMSE statistics. (4) If forecasting all components at different horizons based on a single model is the aim then the 'Average RMSE' statistics can provide the most apt model. (5) Those interested in using univariate modelling can easily determine which SSA model is best given the aim of the analysis.

Table 3: Out-of-sample RMSE for UK IP.

| Variables | $h$ | SSA-R | SSA-V | HMSSA-R | HMSSA-V | $\frac{HMSSA-R}{SSA-R}$ | $\frac{HMSSA-V}{SSA-V}$ |
|---|---|---|---|---|---|---|---|
| Basic Metals | 1 | 6.74 | 6.67 | **6.03**$^R$ | 6.26$^V$ | 0.89 | 0.94 |
| (Fabricated Metals) | 3 | 7.27 | 7.61 | **6.57**$^R$ | 6.78$^V$ | 0.90 | 0.89 |
| | 6 | 7.75 | **6.82**$^V$ | 7.58$^R$ | 7.34 | 0.98 | 1.08 |
| | 12 | 7.52 | 6.83$^V$ | **5.95**$^R$ | 8.44 | 0.79* | 1.24 |
| Fabricated Metals | 1 | 3.84 | 3.72 | **3.58**$^R$ | 3.57$^V$ | 0.93 | 0.96 |
| (Food Products) | 3 | 4.71 | 4.38 | 4.32$^R$ | **4.22**$^V$ | 0.92 | 0.96 |
| | 6 | 4.40 | 4.54 | **4.04**$^R$ | *4.22*$^V$ | 0.92 | 0.93 |
| | 12 | 4.29 | **3.72**$^V$ | **3.72**$^R$ | 3.87 | 0.87 | 1.04 |
| Chemicals | 1 | 2.98 | 3.10 | 2.81$^R$ | **2.74**$^V$ | 0.94 | 0.88 |
| (Electricity and Gas) | 3 | 3.89 | 4.15 | 3.61$^R$ | **3.48**$^V$ | 0.93 | 0.84 |
| | 6 | 4.37 | 4.76 | ***3.78***$^R$ | 4.05$^V$ | 0.86 | 0.85 |
| | 12 | 3.09$^R$ | 3.32 | *3.46* | **3.06**$^V$ | 1.12 | 0.92 |
| Electrical Machinery | 1 | 5.34 | 5.10 | 4.38$^R$ | ***3.98***$^V$ | 0.82 | 0.78* |
| (Vehicles) | 3 | 6.44 | 6.80 | **4.05**$^R$ | 4.75$^V$ | 0.63* | 0.70 |
| | 6 | 5.52$^R$ | 6.10 | *5.60* | ***5.40***$^V$ | 1.01* | 0.89 |
| | 12 | **4.80**$^R$ | 5.31 | *5.49* | *5.09*$^V$ | 1.14* | 0.96 |
| Electricity and Gas | 1 | 4.11 | **3.89**$^V$ | *4.04*$^R$ | *3.98* | 0.98 | 1.02 |
| (Chemicals) | 3 | 4.74 | **4.43** | *4.47*$^R$ | **4.37**$^V$ | 0.94 | 0.99 |
| | 6 | 4.75 | 4.33 | **4.37**$^R$ | *4.24*$^V$ | 0.92* | 0.98 |
| | 12 | 3.74 | 3.06 | ***2.86***$^R$ | *3.35*$^V$ | 0.76* | 1.09* |
| Machinery | 1 | 5.44 | 5.96 | **4.83**$^R$ | 4.97$^V$ | 0.89 | 0.83 |
| (Electrical Machinery) | 3 | 6.56 | 7.33 | 5.82$^R$ | **5.47**$^V$ | 0.89 | 0.75 |
| | 6 | 7.20 | **6.43**$^V$ | 6.76$^R$ | 6.69 | 0.94* | 1.04 |
| | 12 | 5.29 | 5.54$^V$ | **4.96**$^R$ | 6.11 | 0.94 | 1.10 |
| Food Products | 1 | 2.90 | 2.61 | 2.71$^R$ | **2.55**$^V$ | 0.93 | 0.98 |
| (Chemicals) | 3 | **2.39**$^R$ | 2.52$^V$ | *3.15* | 2.85 | 1.32 | 1.13 |
| | 6 | 2.65$^R$ | **2.59**$^V$ | *2.73* | *2.62* | 1.03 | 1.01 |
| | 12 | 2.86 | 2.79 | **2.13**$^R$ | 2.31$^V$ | 0.74* | 0.83* |
| Vehicles | 1 | 8.27 | 8.43 | **7.88**$^R$ | 8.04$^V$ | 0.95 | 0.95 |
| (Electrical Machinery) | 3 | 8.77 | 8.92 | 8.18$^R$ | **8.00**$^V$ | 0.93 | 0.90 |
| | 6 | 9.64 | 9.35 | 9.18$^R$ | ***8.56***$^V$ | 0.95 | 0.92 |
| | 12 | 9.46 | 10.33 | 9.11$^R$ | **9.02**$^V$ | 0.96 | 0.87* |
| Average RMSE | 1 | 4.95 | 4.94 | 4.52 | 4.47 | 0.91 | 0.91 |
| | 3 | 5.60 | 5.76 | 5.02 | 4.99 | 0.90 | 0.87 |
| | 6 | 5.79 | 5.62 | 5.48 | 5.40 | 0.95 | 0.96 |
| | 12 | 5.13 | 5.11 | 4.71 | 5.16 | 0.92 | 1.01 |
| Total RMSE | | 171.72 | 171.38 | 158.15 | 160.38 | - | - |
| Overall Average | | 5.37 | 5.36 | 4.94 | 5.01 | 0.92 | 0.94 |

| | Score | Percentage |
|---|---|---|
| **MSSA>SSA** | **26** | **81%** |
| HMSSA-R>SSA-R | 27 | 84% |
| HMSSA-V>SSA-V | 25 | 78% |
| **SSA>MSSA** | **6** | **19%** |
| SSA-R>HMSSA-R | 5 | 16% |
| SSA-V>HMSSA-V | 7 | 22% |
| HMSSA-R>HMSSA-V | 14 | 44% |
| HMSSA-V>HMSSA-R | 18 | 56% |
| SSA-R>SSA-V | 16 | 50% |
| SSA-V>SSA-R | 16 | 50% |

*Note*: Bold face fonts present the best performing model. Shown in italic are results at lag 12. $^R$ and $^V$ shows the best model using the recurrent or vector approach in comparison to SSA and MSSA. The series shown in () is only applicable to the MSSA forecasts and shows the leading indicator. * indicates a statistically significant difference between the forecasts based on the modified Diebold-Mariano test at $p$=0.10 and a statistically significant difference between the distribution of forecasts based on the two-sided Hassani-Silva (HS) test (Hassani and Silva, 2015) at $p$=0.10.

Table 4: Out-of-sample RMSE for German IP.

| Variables | $h$ | SSA-R | SSA-V | HMSSA-R | HMSSA-V | $\frac{HMSSA-R}{SSA-R}$ | $\frac{HMSSA-V}{SSA-V}$ |
|---|---|---|---|---|---|---|---|
| Basic Metals | 1 | 3.44 | 3.43 | **3.08**$^{R}$ | 3.22$^{V}$ | 0.89 | 0.94 |
| (Electricity and Gas) | 3 | 4.28 | 4.15 | **3.38**$^{R}$ | 3.40$^{V}$ | 0.79 | 0.82 |
| | 6 | 5.42 | 4.60 | ***3.72***$^{R}$ | *3.87*$^{V}$ | 0.69 | 0.84 |
| | 12 | 3.57$^{R}$ | 4.37 | 4.27 | ***3.40***$^{V}$ | 1.20* | 0.78 |
| Fabricated Metals | 1 | 4.55 | 4.38 | ***3.37***$^{R}$ | *4.05*$^{V}$ | 0.74* | 0.92 |
| (Food Products) | 3 | 4.69 | 4.24 | **3.77**$^{R}$ | *4.02*$^{V}$ | 0.80 | 0.94 |
| | 6 | 4.75 | 4.41 | *4.70*$^{R}$ | ***3.90***$^{V}$ | 0.99 | 0.88* |
| | 12 | 4.84 | 4.71 | 4.24$^{R}$ | **4.14**$^{V}$ | 0.88 | 0.88* |
| Chemicals | 1 | 2.81 | 2.58 | 2.62$^{R}$ | **2.50**$^{V}$ | 0.93 | 0.97 |
| (Electricity and Gas) | 3 | 3.00 | 2.57 | *2.61*$^{R}$ | **2.27**$^{V}$ | 0.87 | 0.88* |
| | 6 | 3.52 | 3.05 | 3.04$^{R}$ | **2.61**$^{V}$ | 0.86 | 0.86* |
| | 12 | **2.37**$^{R}$ | 2.53$^{V}$ | *2.58* | 2.78 | 1.09 | 1.10 |
| Electrical Machinery | 1 | 5.09 | **4.95**$^{V}$ | 4.97$^{R}$ | 5.13 | 0.98 | 1.04 |
| (Vehicles) | 3 | 5.21 | 4.74 | 4.87$^{R}$ | **4.63**$^{V}$ | 0.93 | 0.98 |
| | 6 | 5.37 | 4.66 | 3.82$^{R}$ | **3.67**$^{V}$ | 0.71 | 0.79 |
| | 12 | 4.30 | 4.16$^{V}$ | **3.16**$^{R}$ | *4.79* | 0.73* | 1.15 |
| Electricity and Gas | 1 | **3.47**$^{R}$ | 3.55 | 3.56 | 3.53$^{V}$ | 1.03 | 0.99 |
| (Chemicals) | 3 | 3.98 | 3.52 | *3.84*$^{R}$ | **3.51**$^{V}$ | 0.96 | 0.99 |
| | 6 | **3.03**$^{R}$ | 3.71$^{V}$ | 3.16 | 3.73 | 1.04* | 1.01 |
| | 12 | 2.71 | 3.04$^{V}$ | ***2.37***$^{R}$ | 3.16 | 0.87* | 1.04* |
| Machinery | 1 | 5.36 | 5.53 | **4.74**$^{R}$ | 5.14$^{V}$ | 0.88 | 0.93 |
| (Food Products) | 3 | 5.73 | 5.29 | *5.60*$^{R}$ | ***4.41***$^{V}$ | 0.98 | 0.83* |
| | 6 | 6.63$^{R}$ | 6.49 | 6.77 | ***5.71***$^{V}$ | 1.02 | 0.88 |
| | 12 | 5.80 | 6.85 | ***5.20***$^{R}$ | *5.73*$^{V}$ | 0.90* | 0.84* |
| Food Products | 1 | 2.61$^{R}$ | **2.40**$^{V}$ | *2.63* | 2.60 | 1.01 | 1.08 |
| (Machinery) | 3 | 2.61$^{R}$ | 2.55 | 2.71 | **2.49**$^{V}$ | 1.04 | 0.98 |
| | 6 | 2.72 | 2.63 | *2.69*$^{R}$ | **2.42**$^{V}$ | 0.99 | 0.92 |
| | 12 | 2.83$^{R}$ | 2.86 | 3.20 | ***2.47***$^{V}$ | 1.13 | 1.07* |
| Vehicles | 1 | 7.50 | 7.30$^{V}$ | 7.26$^{R}$ | ***7.25*** | 0.97 | 0.99 |
| (Electrical Machinery) | 3 | 7.72 | 7.57 | 7.63$^{R}$ | **7.31**$^{V}$ | 0.99 | 0.97 |
| | 6 | 7.66$^{R}$ | 8.05 | 7.81 | ***7.29***$^{V}$ | 1.02 | 0.91 |
| | 12 | 6.98 | **5.97**$^{V}$ | *6.74*$^{R}$ | *6.21* | 0.97 | 1.04* |
| Average RMSE | 1 | 4.35 | 4.27 | 4.02 | 4.15 | 0.92 | 0.97 |
| | 3 | 4.65 | 4.33 | 4.29 | 4.01 | 0.92 | 0.93 |
| | 6 | 4.89 | 4.70 | 4.41 | 4.15 | 0.90 | 0.88 |
| | 12 | 4.18 | 4.31 | 3.97 | 4.09 | 0.95 | 0.95 |
| Total RMSE | | 144.55 | 140.84 | 133.50 | 131.11 | - | - |
| Overall Average | | 4.52 | 4.40 | 4.17 | 4.10 | 0.92 | 0.93 |

| | Score | Percentage |
|---|---|---|
| **MSSA>SSA** | **26** | **81%** |
| HMSSA-R>SSA-R | 23 | 72% |
| HMSSA-V>SSA-V | 25 | 78% |
| **SSA>MSSA** | **6** | **19%** |
| SSA-R>HMSSA-R | 9 | 28% |
| SSA-V>HMSSA-V | 7 | 22% |
| HMSSA-R>HMSSA-V | 12 | 38% |
| HMSSA-V>HMSSA-R | 20 | 62% |
| SSA-R>SSA-V | 9 | 28% |
| SSA-V>SSA-R | 23 | 72% |

*Note*: Bold face fonts present the best performing model. Shown in italic are results at lag 12. $^{R}$ and $^{V}$ shows the best model using the recurrent or vector approach in comparison to SSA and MSSA. The series shown in () is only applicable to the MSSA forecasts and shows the leading indicator. * indicates a statistically significant difference between the forecasts based on the modified Diebold-Mariano test at $p$=0.10 and a statistically significant difference between the distribution of forecasts based on the two-sided HS test at $p$=0.10.

Table 5: Out-of-sample RMSE for France IP.

| Variables | $h$ | SSA-R | SSA-V | HMSSA-R | HMSSA-V | $\frac{HMSSA-R}{SSA-R}$ | $\frac{HMSSA-V}{SSA-V}$ |
|---|---|---|---|---|---|---|---|
| Basic Metals | 1 | 4.64 | 4.58 | 4.28$^R$ | **4.03**$^V$ | 0.92 | 0.88 |
| (Food Products) | 3 | 5.16 | 4.83 | 4.54$^R$ | **4.18**$^V$ | 0.88 | 0.87* |
| | 6 | 3.73 | 4.11 | **3.21**$^R$ | 3.98$^V$ | 0.86 | 0.97* |
| | 12 | **3.94**$^R$ | 4.34$^V$ | 4.26 | 4.45 | 1.08* | 1.03* |
| Fabricated Metals | 1 | 3.00 | **2.64**$^V$ | *2.68*$^R$ | 2.67 | 0.89 | 1.01 |
| (Electricity & Gas) | 3 | 3.37 | 2.87$^V$ | **2.83**$^R$ | 2.88 | 0.84 | 1.00 |
| | 6 | 2.64 | 2.93 | 2.51$^R$ | **2.42**$^V$ | 0.95 | 0.83 |
| | 12 | 3.08 | 2.75 | 2.84$^R$ | **2.42**$^V$ | 0.92 | 0.88* |
| Chemicals | 1 | 4.36 | 3.94 | **3.64**$^R$ | 3.70$^V$ | 0.83 | 0.94 |
| (Electricity & Gas) | 3 | 5.48 | 4.19 | ***3.38***$^R$ | 4.12$^V$ | 0.62* | 0.98* |
| | 6 | 4.15 | 3.77 | 3.63$^R$ | ***3.31***$^V$ | 0.87 | 0.88 |
| | 12 | 3.58 | 3.32 | *3.03*$^R$ | ***2.80***$^V$ | 0.85* | 0.84 |
| Electrical & Electronic | 1 | 4.02 | 4.06 | ***3.58***$^R$ | 3.63$^V$ | 0.89 | 0.89 |
| (Basic Metals) | 3 | 4.31 | 4.09 | ***3.40***$^R$ | 3.90$^V$ | 0.79 | 0.95 |
| | 6 | 3.37 | 4.05 | **3.10**$^R$ | 3.75$^V$ | 0.92 | 0.93 |
| | 12 | 4.03 | 3.42 | 3.75$^R$ | ***3.34***$^V$ | 0.93* | 0.98 |
| Electricity & Gas | 1 | 6.16 | 6.08 | ***5.64***$^R$ | *5.72*$^V$ | 0.92 | 0.94 |
| (Basic Metals) | 3 | 6.71 | 6.38 | *5.84*$^R$ | **5.82**$^V$ | 0.87 | 0.91 |
| | 6 | 6.23 | 6.18 | 5.91$^R$ | **5.55**$^V$ | 0.95 | 0.90 |
| | 12 | 5.12 | 4.18 | ***3.37***$^R$ | 3.98$^V$ | 0.66* | 0.95 |
| Machinery | 1 | 6.16 | 6.11 | **5.51**$^R$ | 5.77$^V$ | 0.89 | 0.94 |
| (Vehicles) | 3 | 6.38 | 7.05 | **5.48**$^R$ | 5.88$^V$ | 0.86 | 0.83 |
| | 6 | 5.13 | 6.36 | **4.67**$^R$ | *5.88*$^V$ | 0.91 | 0.92 |
| | 12 | 5.53 | 5.74 | 4.87$^R$ | **4.65**$^V$ | 0.88 | 0.81* |
| Food Products | 1 | **1.80**$^R$ | 1.81$^V$ | 2.02 | *1.85* | 1.12 | 1.02 |
| (Electricity & Gas) | 3 | **1.82**$^R$ | 1.83$^V$ | 2.00 | 1.88 | 1.10 | 1.03 |
| | 6 | 1.78$^R$ | **1.75**$^V$ | 1.86 | 1.83 | 1.04 | 1.05 |
| | 12 | **1.61**$^R$ | 1.69$^V$ | 1.69 | 1.72 | 1.05 | 1.02* |
| Vehicles | 1 | 7.57 | 7.58 | **7.47**$^R$ | 7.48$^V$ | 0.99 | 0.99 |
| (Fabricated Metals) | 3 | 7.92 | 8.09 | *7.76*$^R$ | **7.67**$^V$ | 0.98 | 0.95 |
| | 6 | 7.49 | 7.67 | **6.90**$^R$ | 7.21$^V$ | 0.92 | 0.94 |
| | 12 | 9.83 | 7.90 | **7.71**$^R$ | *7.80*$^V$ | 0.78* | 0.99 |
| Average RMSE | 1 | 4.71 | 4.60 | 4.36 | 4.35 | 0.92 | 0.95 |
| | 3 | 5.14 | 4.92 | 4.40 | 4.54 | 0.86 | 0.92 |
| | 6 | 4.32 | 4.60 | 3.97 | 4.24 | 0.92 | 0.92 |
| | 12 | 4.59 | 4.17 | 3.90 | 3.90 | 0.85 | 0.93 |
| Total RMSE | | 150.1 | 146.29 | 133.09 | 136.18 | - | - |
| Overall Average | | 4.69 | 4.57 | 4.16 | 4.26 | 0.89 | 0.93 |

| | Score | Percentage |
|---|---|---|
| **MSSA>SSA** | **26** | **81%** |
| HMSSA-R>SSA-R | 27 | 84% |
| HMSSA-V>SSA-V | 25 | 78% |
| **SSA>MSSA** | **6** | **19%** |
| SSA-R>HMSSA-R | 5 | 16% |
| SSA-V>HMSSA-V | 7 | 22% |
| HMSSA-R>HMSSA-V | 17 | 53% |
| HMSSA-V>HMSSA-R | 15 | 47% |
| SSA-R>SSA-V | 14 | 44% |
| SSA-V>SSA-R | 18 | 56% |

*Note*: Bold face fonts present the best performing model. Shown in italic are results at lag 12. $^R$ and $^V$ shows the best model using the recurrent or vector approach in comparison to SSA and MSSA. The series shown in () is only applicable to the MSSA forecasts and shows the leading indicator. * indicates a statistically significant difference between the forecasts based on the modified Diebold-Mariano test at $p$=0.10 and a statistically significant difference between the distribution of forecasts based on the two-sided HS test at $p$=0.10.

## 5 Conclusion

The aim of this paper was to introduce an optimized MSSA forecasting algorithm for detecting leading indicators in European industrial production. The importance of this algorithm is not only for speeding up the overall MSSA process, but also for enabling users who lack an in-depth knowledge of the methodology underlying MSSA to exploit this powerful, nonparametric time series analysis and forecasting technique in practice. It is expected that algorithms such as the one introduced in this paper will encourage increased applications of MSSA in the future. The idea of optimizing the MSSA algorithm is based on the work in (Hassani et al., 2015) where such an algorithm was introduced for univariate SSA. The proposed approach saves time and effort by enabling users to obtain the best possible SSA choices $(L, r)$ which minimises a loss function for any given data set.

Using forecasts from the optimized univariate SSA algorithm (Hassani et al., 2015) as a benchmark, we forecast 24 monthly, seasonally unadjusted components of industrial production for UK, Germany and France with optimized HMSSA-R and HMSSA-V algorithms. Prior to the forecasting exercise, the data are tested for seasonal unit root problems (using the OCSB test in Osborn et al., (1988)) and structural breaks (using the Bai and Perron breakpoints test in Bai and Perron (2003)). The out-of-sample forecasting errors are evaluated via the RMSE and RRMSE criterions with detailed results which consider the score, percentage, total and average RMSE/RRMSE for presenting conclusions based on our findings. The forecasts themselves are evaluated for statistically significant differences using both the modified DM test (Harvey et al., 1997) and the HS test (Hassani and Silva, 2015).

In general, we find the proposed MSSA algorithms outperforming the SSA algorithms and proving that they are indeed viable. However, it is noteworthy that no single model (either SSA or MSSA) can provide the best forecast for each component in any given country across all horizons. For UK, based on the score we have the following to report. In comparison to its univariate counterpart, HMSSA-R provides the best performance with the highest score. However, when compared with HMSSA-V, the latter reports the highest score. Based on the overall lowest average RMSE, HMSSA-R is more likely to provide more accurate forecasts than HMSSA-V. Based on the RRMSE, HMSSA-R forecasts are 8% better than SSA-R forecasts whilst HMSSA-V forecasts are 6% better than SSA-V forecasts. For Germany, based on the score, HMSSA-V not only reports the best performance in comparison to SSA-V, but also in comparison to HMSSA-R. Based on the lowest average RMSE, both HMSSA-R and HMSSA-V are almost identical. The RRMSE shows that HMSSA-R forecasts continue to be 8% better than SSA-R forecasts whilst the HMSSA-V forecasts are found to be 7% better than SSA-V forecasts for Germany. For France, based on the score we can conclude that HMSSA-R reports the best performance in comparison to its univariate counterpart and in comparison to HMSSA-V. This is also confirmed by the overall lowest average RMSE whereby, HMSSA-R forecasts are seen outperforming HMSSA-V forecasts. The RRMSE indicates that HMSSA-R forecasts are 11% better than SSA-R and that HMSSA-V forecasts are 7% better than SSA-V forecasts. For those interested in univariate models, based on the score for both Germany and France we notice that SSA-R outperforms SSA-V whilst in UK the two models are indifferent as they share the same score.

This paper opens up several avenues for future research in this area. Firstly, it would be interesting to ascertain the performance of other time series analysis models, both univariate and multivariate (which have not been compared in previous literature using this same data) in comparison to the proposed MSSA algorithm when forecasting IP. Secondly, it would be interesting to evaluate the outcome if a mixed approach was used whereby one considers both univariate and multivariate SSA models for forecasting IP in these countries and compare the results alongside other competing models. Also, it is noteworthy that the proposed optimized MSSA algorithm is not without its drawbacks. It is important to remember that this algorithm is optimized for obtaining the best possible out-of-sample forecasts. It is not optimized for the best MSSA decomposition or for the best possible signal extraction in a multivariate framework. Finally, future work could also concentrate on evaluating different cross country analysis to find leading indicators across countries for industrial production.

## References


[1] Altavilla, C., and De Grauwe, P. (2010). Forecasting and combining competing models of exchange rate determination. *Applied Economics*, **42**, 345580.

[2] Bai J., and Perron P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, **18**(1), 1–22.

[3] Broomhead, D. S., and King, G. P. (1986a). Extracting Qualitative Dynamics from Experimental Data. *Physica D: Nonlinear Phenomena*, **20**(2-3), 217–236.

[4] Broomhead, D. S., and King, G. P. (1986b). On the Qualitative Analysis of Experimental Dynamical Systems. *Nonlinear Phenomena and Chaos*, ed. S. Sarkar, Adam Hilger, Bristol.

[5] de Carvalho, M., and Rua, A. (2017). Real-time nowcasting the US output gap: Singular spectrum analysis at work. *International Journal of Forecasting*, **33**(1), 185–198.

[6] Ghodsi, Z., Silva, E. S., and Hassani, H. (2015). *Bicoid* Signal Extraction with a Selection of Parametric and Nonparametric Signal Processing Techniques. *Genomics Proteomics & Bioinformatics*, **13**, 183–191.

[7] Gong, Y., Song, Z., He, M., Gong, W., and Ren, F. (2017). Precursory waves and eigenfrequencies identified from acoustic emission data based on Singular Spectrum Analysis and laboratory rock-burst experiments. *International Journal of Rock Mechanics and Mining Sciences*, **91**, 155–169.

[8] Hassani, H., Silva, E. S., Antonakakis, N., Filis, G., and Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research*, **63**, 112–127.

[9] Hassani, H., and Silva, E. S. (2016). Forecasting energy data with a time lag into the future and Google trends. *International Journal of Energy and Statistics*, **4**(4), 1650020.

[10] Hassani, H., Ghodsi, Z., Silva, E. S., and Heravi, S. (2016). From nature to maths: Improving forecasting performance in subspace-based methods using genetics Colonial Theory. *Digital Signal Processing*, **51**, 101–109.

[11] Hassani, H., and Silva, E. S. (2015). A Kolmogorov-Smirnov Based Test for Comparing the Predictive Accuracy of Two Sets of Forecasts. *Econometrics*, **3**(3), 590–609.

[12] Hassani, H., Webster, A., Silva, E. S., and Heravi, S. (2015). Forecasting U.S. Tourist Arrivals using Optimal Singular Spectrum Analysis. *Tourism Management*, **46**, 322–335.

[13] Hassani, H., Heravi, H., and Zhigljavsky, A. (2013). Forecasting UK industrial production with multivariate singular spectrum analysis. *Journal of Forecasting*, **32**(5), 395–408.

[14] Hassani, H., Heravi, H., and Zhigljavsky, A. (2009). Forecasting European industrial production with Singular Spectrum Analysis. *International Journal of Forecasting*, **25**(1), 103–118.

[15] Heravi, S., Osborn, D. R., and Birchenhall, C. R. (2004). Linear versus neural network forecasts for European industrial production series. *International Journal of Forecasting*, **20**(3), 435–446.

[16] Hirashima, A., Jones, J., Bonham, C. S., and Fuleky, P. (2017). Forecasting in a Mixed Up World: Nowcasting Hawaii Tourism. *Annals of Tourism Research*, **63**, 191–202.

[17] Khan, M. A. R., and Poskitt, D. (2017). Forecasting stochastic processes using singular spectrum analysis: Aspects of the theory and application. *International Journal of Forecasting*, **33**(1), 199–213.

[18] Khan, M. A. R., and Poskitt, D. (2013). A note on window length selection in Singular Spectrum Analysis. *Australian and New Zealand Journal of Statistics*, **55**(2), 87–108.

[19] Lai, L., and Guo, K. (2017). The performance of one belt and one road exchange rate: Based on improved singular spectrum analysis. *Physica A: Statistical Mechanics and its Applications*, **483**, 299–308.

[20] Mahmoudvand, R., and Rodrigues, P. C. (2017). A new parsimonious recurrent forecasting model in singular spectrum analysis. *Journal of Forecasting*, DOI: 10.1002/for.2484.

[21] Merte, S. (2017). Estimating heat wave-related mortality in Europe using singular spectrum analysis. *Climatic Change*, **142**(3-4), 321–330.

[22] Miron, J. A. (1996). *The economics of seasonal cycles*. Cambridge, MA: MIT Press.

[23] Osborn, D. R., Heravi, S., and Birchenhall, C. R. (1999). Seasonal unit roots and forecasts of two-digit European industrial production. *International Journal of Forecasting*, **15**(1), 27–47.

[24] Osborn, D. R., Chui, A. P. L., Smith, J., and Birchenhall, C. R. (1988). Seasonality and the order of integration for consumption. *Oxford Bulletin of Economics and Statistics*, **50**(4), 361–377.

[25] Portes, L. L., and Aguirre, L. A. (2016). Enhancing multivariate singular spectrum analysis for phase synchronization: The role of observability. *Chaos*, **26**, 093112.

[26] Royston, P. (1982a). An extension of Shapiro and Wilk's $W$ test for normality to large samples. *Applied Statistics*, **31**(2), 115–124.

[27] Royston, P. (1982b). Algorithm AS 181: The $W$ test for Normality. *Applied Statistics*, **31**(2), 176–180.

[28] Royston, P. (1995). Remark AS R94: A remark on Algorithm AS 181: The $W$ test for normality. *Applied Statistics*, **44**(4), 547–551.

[29] Sadorsky, P. (2016). Forecasting Canadian mortgage rates. *Applied Economic Letters*, **23**(11), 822–825.

[30] Sanei, S., and Hassani, H. (2015). *Singular spectrum analysis of biomedical signals*. United States: CRC Press.

[31] Silva, E. S. and Hassani, H. (2015). On the use of Singular Spectrum Analysis for Forecasting U.S. Trade before, during and after the 2008 Recession. *International Economics*, doi:10.1016/j.inteco.2014.11.003.

[32] Silva, E .S., Ghodsi, Z., Ghodsi, M., Heravi, S., and Hassani, H. (2017). Cross country relations in European tourist arrivals. *Annals of Tourism Research*, **63**, 151–168.

[33] Yu, C., Li, Y., and Zhang, M. (2017). An improved Wavelet Transform using Singular Spectrum Analysis for wind speed forecasting based on Elman Neural Network. *Energy Conversion and Management*, **148**, 895–904.